UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/714,465	11/14/2003	Jinquan Dai	42P17829	2488

8791 7590 07/06/2007

BLAKELY SOKOLOFF TAYLOR & ZAFMAN
1279 OAKMEAD PARKWAY
SUNNYVALE, CA 94085-4040

EXAMINER	
WANG, BEN C	
ART UNIT	**PAPER NUMBER**
2192	
MAIL DATE	**DELIVERY MODE**
07/06/2007	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No.	Applicant(s)	
	10/714,465	DAI ET AL.	
	Examiner	Art Unit	
	Ben C. Wang	2192	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *14 November 2003*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-36* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-4,8-14,18-21,26 and 31-36* is/are rejected.

7)☒ Claim(s) *5-7,15-17,22-25 and 27-30* is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date *11/14/2003, 8/8/2005*.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. Claims 1-36 are pending in this application and presented for examination.

Specification Objections

2. The specification is objected to because the following informalities:

- "As depicted in FIG. 2B, PPS 280 may be", cited in [00029], Line 2, should be corrected as "As depicted in FIG. 2A, PPS 280 may be",
- "a D-stage processor pipeline of network processor 500 of FIG. 1", cited in [00029], Lines 4-5, should be corrected as "a D-stage processor pipeline of network processor 500 of FIG. 8"

Appropriate correction is required

Claim Objections

3. Claims 5, 10, 15, 20, and 21 are objected to because the following informalities:

- "comprises;", cited in claim 5, line 3, should be corrected as "comprises:"
- "altering the selected preliminary pipeline stage to enable proper transmission of live variables to and from the selected preliminary pipeline stage; altering the selected preliminary pipeline stage to enable proper transmission of control flow to and from the selected preliminary pipeline stage;", cited in claim 10, lines 4-7; claim 20, lines 4-7 respectively, should be corrected as "altering the selected preliminary pipeline stage to enable proper transmission of live variables and control flow to and from the selected preliminary pipeline

stage;"; "repeating the selecting, altering and altering for each preliminary pipeline stage", cited in claim 10, line 8; claim 20, line 8 respectively, should be corrected as "repeating the above 'selecting' and 'altering' steps for each preliminary pipeline stage"

- "comprises;", cited in claim 15, line 3, should be corrected as "comprises:"
- "comprising;", cited in claim 21, line 1, should be corrected as "comprising:"

Appropriate correction is required.

Claim Rejections – 35 USC § 102(b)

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102(b) that form the basis for the rejections under this section made in this office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

4. Claims 1, 11, 21, 26, 31, and 34 are rejected under 35 U.S.C. 102(b) as being anticipated by Lakshmanamurthy et al. (*"Network Processor Performance Analysis Methodology", Aug. 15, 2002, Intel Technology Journal*) (hereinafter 'Lakshmanamurthy')

5. **As to claim 1**, Lakshmanamurthy discloses a method comprising configuring one or more processors into a D-stage processor pipeline; transforming a sequential application program into D-pipeline stages; and executing the D-pipeline stages in

parallel within the D-stage processor pipeline (Sec. of "ABSTRACT", 1st Par. – this paper describes the performance analysis methodology developed to analyze the performance of various networking applications that are targeted for running on the IXP 2400 network processor, the second-generation IXA network processor; P. 19, L-Col., 3rd Par., Lines 4-9 – this methodology involves diving the application into pipeline blocks.. and latency budget for each pipeline element, and mapping the application blocks to software paradigms and the hardware resources; P. 20, L-Col., Lines 9-12 – how the software concepts (hyper-task chaining,, pool of threads, functional pipeline, context pipeline) are used to meet the performance goals; P. 21, L-Col., Lines 8-13 – IXP 2400 also offers extensive communication mechanisms between all on-chip processing units and enables the micro-engines to readily form different topologies of software pipelines that can be customized for various target applications and network traffic patterns; P. 21, R-Col., 3rd Par., Lines 8-11 – these innovations enable the micro-engines to form various topologies of software pipelines flexibly and efficiently, allowing processing to be tuned to specific applications and traffic patterns; P. 22, L-Col., 4th Par., Lines 3-7 – in order to keep up with arrival rate, no processing stage in the pipeline can exceed this inter-arrival time; P. 22, R-Col., 6th Par., Lines 3-6 – in order to keep up with the arrival rate, any processing stage of the pipeline must complete all the required processing for a given packet....; P. 23, L-Col., 3rd Par. Through R-Col., 3rd Par.; P. 24, L-Col., 4th Par. Through R-Col., 2nd Par.; P. 24, Sec. of "Performance Analysis Methodology"; P. 25, Sec of "Classification Pipeline"; Sec. of "CONCLUSION" – this methodology involves estimation of the available processing and latency budget per

packet, estimation of the total compute and IO operations required per packet for the various pipeline stages of the target application, and mapping the pipeline stages of the application to the available hardware resources on the IXP 2400 and utilizing all the available software pipelining techniques to hit the desired performance).

6. **As to claim 11**, Lakshmanamurthy discloses an article of manufacture including a machine readable medium having stored thereon instructions which may be used to program a system to perform a method, comprising configuring one or more processors into a D-stage processor pipeline; transforming a sequential application program into D-pipeline stages; and executing the D-pipeline stages in parallel within the D-stage processor pipeline (Sec. of "ABSTRACT", 1st Par. – this paper describes the performance analysis methodology developed to analyze the performance of various networking applications that are targeted for running on the IXP 2400 network processor, the second-generation IXA network processor; P. 19, L-Col., 3rd Par., Lines 4-9 – this methodology involves diving the application into pipeline blocks.. and latency budget for each pipeline element, and mapping the application blocks to software paradigms and the hardware resources; P. 20, L-Col., Lines 9-12 – how the software concepts (hyper-task chaining,, pool of threads, functional pipeline, context pipeline) are used to meet the performance goals; P. 21, L-Col., Lines 8-13 – IXP 2400 also offers extensive communication mechanisms between all on-chip processing units and enables the micro-engines to readily form different topologies of software pipelines that can be customized for various target applications and network traffic patterns; P. 21, R-Col., 3rd Par., Lines 8-11 – these innovations enable the micro-engines to form various

topologies of software pipelines flexibly and efficiently, allowing processing to be tuned to specific applications and traffic patterns; P. 22, L-Col., 4th Par., Lines 3-7 – in order to keep up with arrival rate, no processing stage in the pipeline can exceed this inter-arrival time; P. 22, R-Col., 6th Par., Lines 3-6 – in order to keep up with the arrival rate, any processing stage of the pipeline must complete all the required processing for a given packet....; P. 23, L-Col., 3rd Par. Through R-Col., 3rd Par.; P. 24, L-Col., 4th Par. Through R-Col., 2nd Par.; P. 24, Sec. of "Performance Analysis Methodology"; P. 25, Sec of "Classification Pipeline"; Sec. of "CONCLUSION" – this methodology involves estimation of the available processing and latency budget per packet, estimation of the total compute and IO operations required per packet for the various pipeline stages of the target application, and mapping the pipeline stages of the application to the available hardware resources on the IXP 2400 and utilizing all the available software pipelining techniques to hit the desired performance).

7. **As to claim 21**, Lakshmanamurthy discloses a method comprising constructing a flow network model from a sequential application program; cutting the flow network model into a plurality of preliminary pipeline stages; and transforming the preliminary pipeline stages to perform control flow and variable transmission therebetween in order to form D-pipeline stages of a parallel application program (Sec. of "ABSTRACT", 1st Par. – this paper describes the performance analysis methodology developed to analyze the performance of various networking applications that are targeted for running on the IXP 2400 network processor, the second-generation IXA network processor; P.

19, L-Col., 3rd Par., Lines 4-9 – this methodology involves diving the application into pipeline blocks.. and latency budget for each pipeline element, and mapping the application blocks to software paradigms and the hardware resources; P. 20, L-Col., Lines 9-12 – how the software concepts (hyper-task chaining,, pool of threads, functional pipeline, context pipeline) are used to meet the performance goals; P. 19, R-Col., 3rd Par. Through Col. 20, L-Col., 2nd Par.; P. 21, L-Col., Lines 8-13 – IXP 2400 also offers extensive communication mechanisms between all on-chip processing units and enables the micro-engines to readily form different topologies of software pipelines that can be customized for various target applications and network traffic patterns; P. 21, R-Col., 3rd Par., Lines 8-11 – these innovations enable the micro-engines to form various topologies of software pipelines flexibly and efficiently, allowing processing to be tuned to specific applications and traffic patterns; P. 22, L-Col., 4th Par., Lines 3-7 – in order to keep up with arrival rate, no processing stage in the pipeline can exceed this inter-arrival time; P. 22, R-Col., 6th Par., Lines 3-6 – in order to keep up with the arrival rate, any processing stage of the pipeline must complete all the required processing for a given packet....; P. 23, L-Col., 3rd Par. Through R-Col., 3rd Par.; P. 24, L-Col., 4th Par. Through R-Col., 2nd Par.; P. 24, Sec. of "Performance Analysis Methodology"; P. 25, Sec of "Classification Pipeline"; Sec. of "CONCLUSION" – this methodology involves estimation of the available processing and latency budget per packet, estimation of the total compute and IO operations required per packet for the various pipeline stages of the target application, and mapping the pipeline stages of the application to the

available hardware resources on the IXP 2400 and utilizing all the available software pipelining techniques to hit the desired performance).

8. **As to claim 26**, Lakshmanamurthy discloses an article of manufacture including a machine readable medium having stored thereon instructions which may be used to program a system to perform a method, comprising constructing a flow network model from a sequential application program; cutting the flow network model into a plurality of preliminary pipeline stages; and transforming the preliminary pipeline stages to perform control flow and variable transmission therebetween in order to form D-pipeline stages of a parallel application program (Sec. of "ABSTRACT", 1st Par. – this paper describes the performance analysis methodology developed to analyze the performance of various networking applications that are targeted for running on the IXP 2400 network processor, the second-generation IXA network processor; P. 19, L-Col., 3rd Par., Lines 4-9 – this methodology involves diving the application into pipeline blocks.. and latency budget for each pipeline element, and mapping the application blocks to software paradigms and the hardware resources; P. 20, L-Col., Lines 9-12 – how the software concepts (hyper-task chaining,, pool of threads, functional pipeline, context pipeline) are used to meet the performance goals; P. 21, L-Col., Lines 8-13 – IXP 2400 also offers extensive communication mechanisms between all on-chip processing units and enables the micro-engines to readily form different topologies of software pipelines that can be customized for various target applications and network traffic patterns; P. 21, R-Col., 3rd Par., Lines 8-11 – these innovations enable the micro-engines to form various

topologies of software pipelines flexibly and efficiently, allowing processing to be tuned to specific applications and traffic patterns; P. 22, L-Col., 4th Par., Lines 3-7 – in order to keep up with arrival rate, no processing stage in the pipeline can exceed this inter-arrival time; P. 22, R-Col., 6th Par., Lines 3-6 – in order to keep up with the arrival rate, any processing stage of the pipeline must complete all the required processing for a given packet....; P. 23, L-Col., 2nd Par. Through R-Col., 3rd Par.; P. 24, L-Col., 4th Par. Through R-Col., 2nd Par.; P. 24, Sec. of "Performance Analysis Methodology"; P. 25, Sec of "Classification Pipeline"; Sec. of "CONCLUSION" – this methodology involves estimation of the available processing and latency budget per packet, estimation of the total compute and IO operations required per packet for the various pipeline stages of the target application, and mapping the pipeline stages of the application to the available hardware resources on the IXP 2400 and utilizing all the available software pipelining techniques to hit the desired performance).

9. **As to claim 31**, Lakshmanamurthy discloses an apparatus, comprising a processor (Sec. of "ABSTRACT", 1st Par. – this paper describes the performance analysis methodology developed to analyze the performance of various networking applications that are targeted for running on the IXP 2400 network processor, the second-generation IXA network processor; Fig. 1 – IXP 2400 external interface, element of "IXP 2400"; P. 20, R-Col., 1st Par.; P. 21, L-Col., 3rd Par. – IXP 2400 contains eight multi-threaded, packet-processing micro-engines; these micro-engines are highly programmable packet processors and support multi threading of up to eight threads

each; each micro-engine provides a variety of network processing functions in hardware; P. 21, R-Col., 2nd Par. – the IXP 2400 also has an integrated low-power general-purpose Intel® Xscale™ micro-architecture core; the integrated Xscale™ process offers ample processing power for running control plane software); a memory coupled to the processor (P. 20, L-Col., 4th Par. – extern DRAM and SRAM; P. 20, L-Col., 4th Par. – P. 21, L-Col., 1st Par. – the SRAM is primarily used for packet descriptors, queue descriptors, counters, and other data structures; Fig. 2 – IXP 2400 internal architecture, elements of "QDR SRAM", "DDRAM"; Fig. 3 – IXP 2400-based OC-48 line card configuration, element of "DDR SDRAM"), the memory including a compiler to cause transformation of a sequential application program into D-pipeline stages to enable parallel execution of the D-pipeline stages within a D-stage processor pipeline (P. 19, L-Col., 3rd Par., Lines 4-9 – this methodology involves diving the application into pipeline blocks.. and latency budget for each pipeline element, and mapping the application blocks to software paradigms and the hardware resources; P. 20, L-Col., Lines 9-12 – how the software concepts (hyper-task chaining,, pool of threads, functional pipeline, context pipeline) are used to meet the performance goals; P. 21, L-Col., Lines 8-13 – IXP 2400 also offers extensive communication mechanisms between all on-chip processing units and enables the micro-engines to readily form different topologies of software pipelines that can be customized for various target applications and network traffic patterns; P. 21, R-Col., 3rd Par., Lines 8-11 – these innovations enable the micro-engines to form various topologies of software pipelines flexibly and efficiently, allowing processing to be tuned to specific applications and traffic

patterns; P. 22, L-Col., 4th Par., Lines 3-7 – in order to keep up with arrival rate, no processing stage in the pipeline can exceed this inter-arrival time; P. 22, R-Col., 6th Par., Lines 3-6 – in order to keep up with the arrival rate, any processing stage of the pipeline must complete all the required processing for a given packet....; P. 23, L-Col., 3rd Par. Through R-Col., 3rd Par.; P. 24, L-Col., 4th Par. Through R-Col., 2nd Par.; P. 24, Sec. of "Performance Analysis Methodology"; P. 25, Sec of "Classification Pipeline"; Sec. of "CONCLUSION" – this methodology involves estimation of the available processing and latency budget per packet, estimation of the total compute and IO operations required per packet for the various pipeline stages of the target application, and mapping the pipeline stages of the application to the available hardware resources on the IXP 2400 and utilizing all the available software pipelining techniques to hit the desired performance).

10. **As to claim 34**, Lakshmanamurthy discloses a system comprising a processor (Sec. of "ABSTRACT", 1st Par. – this paper describes the performance analysis methodology developed to analyze the performance of various networking applications that are targeted for running on the IXP 2400 network processor, the second-generation IXA network processor; Fig. 1 – IXP 2400 external interface, element of "IXP 2400"; P. 20, R-Col., 1st Par.; P. 21, L-Col., 3rd Par. – IXP 2400 contains eight multi-threaded, packet-processing micro-engines; these micro-engines are highly programmable packet processors and support multi threading of up to eight threads each; each micro-engine provides a variety of network processing functions in hardware; P. 21, R-Col., 2nd Par. – the IXP 2400 also has an integrated low-power general-purpose Intel® Xscale™ micro-

architecture core; the integrated Xscale™ process offers ample processing power for running control plane software); a memory controller coupled to the processor (P. 21, L-Col., 3rd Par., Lines 13-14 – the memory controllers facilitate efficient access to the of-chip SRAM and DRAM); and a DDR SRAM memory coupled to the processor (P. 20, L-Col., 4th Par. – extern DRAM and SRAM; P. 20, L-Col., 4th Par. – P. 21, L-Col., 1st Par. – the SRAM is primarily used for packet descriptors, queue descriptors, counters, and other data structures; Fig. 2 – IXP 2400 internal architecture, element of "QDR SRAM"; Fig. 3 – IXP 2400-based OC-48 line card configuration, element of "DDR SDRAM"), the memory including a compiler to cause transformation of a sequential application program into D-application program stages to enable parallel execution of the D-application program stages within a D-stage processor pipeline (P. 19, L-Col., 3rd Par., Lines 4-9 – this methodology involves diving the application into pipeline blocks.. and latency budget for each pipeline element, and mapping the application blocks to software paradigms and the hardware resources; P. 20, L-Col., Lines 9-12 – how the software concepts (hyper-task chaining,, pool of threads, functional pipeline, context pipeline) are used to meet the performance goals; P. 21, L-Col., Lines 8-13 – IXP 2400 also offers extensive communication mechanisms between all on-chip processing units and enables the micro-engines to readily form different topologies of software pipelines that can be customized for various target applications and network traffic patterns; P. 21, R-Col., 3rd Par., Lines 8-11 – these innovations enable the micro-engines to form various topologies of software pipelines flexibly and efficiently, allowing processing to be tuned to specific applications and traffic patterns; P. 22, L-Col., 4th Par., Lines 3-7 – in

order to keep up with arrival rate, no processing stage in the pipeline can exceed this inter-arrival time; P. 22, R-Col., 6th Par., Lines 3-6 – in order to keep up with the arrival rate, any processing stage of the pipeline must complete all the required processing for a given packet....; P. 23, L-Col., 3rd Par. Through R-Col., 3rd Par.; P. 24, L-Col., 4th Par. Through R-Col., 2nd Par.; P. 24, Sec. of "Performance Analysis Methodology"; P. 25, Sec of "Classification Pipeline"; Sec. of "CONCLUSION" – this methodology involves estimation of the available processing and latency budget per packet, estimation of the total compute and IO operations required per packet for the various pipeline stages of the target application, and mapping the pipeline stages of the application to the available hardware resources on the IXP 2400 and utilizing all the available software pipelining techniques to hit the desired performance).

Claim Rejections – 35 USC § 103(a)

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

11. Claims 2, 12, 32, and 35 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lakshmanamurthy in view of Rakhmatov et al., ("*Hardware-Software Bipartitioning for Dynamically Reconfigurable Systems", May 2002, ACM*) (hereinafter 'Rakhmatov')

12. **As to claim 2** (incorporating the rejection in claim 1) and **claim 12** (incorporating the rejection in claim 11), Lakshmanamurthy discloses network processor performance analysis methodology (Sec. of "ABSTRACT", 3rd Par.), but does not explicitly disclose transforming the sequential application program comprises constructing a flow network model for the sequential application program; selecting a plurality of preliminary pipeline stages from the flow network model; and modifying the preliminary pipeline stages to perform control flow and variable transmission therebetween to form the D-pipeline stages.

However, in an analogous art of hardware-software bi-partitioning for dynamically reconfigurable systems, Rakhmatov discloses transforming the sequential application program comprises constructing a flow network model for the sequential application program; selecting a plurality of preliminary pipeline stages from the flow network model; and modifying the preliminary pipeline stages to perform control flow and variable transmission therebetween to form the D-pipeline stages (Sec. of "ABSTRACT" – a method for mapping nodes of an application control flow graph either to software or reconfigurable hardware, explicitly targeting minimization of the energy-delay cost due to both computation and configuration; using network flow techniques, after transforming the original control flow graph into an equivalent network; P. 145, R-Col., 1st Par. Through 3rd Par. – the software can directly configure the hardware, which is partially reconfigurable; partial reconfiguration allows for a selective change of hardware segments of arbitrary size at an arbitrary location, without disrupting the operation of the rest of the hardware space; such a capability greatly reduces reconfiguration time and

energy consumption, because the hardware updates are highly localized. Three types of problems: (1) energy-delay product minimization, (2) energy minimization under the delay constraint, and (3) delay minimization under the energy constraint can be resolved via using network flow techniques; specifically, the cost of a node depends whether it is in software or in hardware and the cost of an edge depends whether its origin node is in software or in hardware and whether its destination node is in software or in hardware; P. 146, L-Col., 2nd Par. – 4th Par. – the cost/weight can be either the energy or the delay or the energy-delay product of a node/edge, weighted by its execution frequency; first, transferring control from a hardware block to a software block is more expensive than transferring control from a software block to a software block; second, transferring control from a software block to a hardware block is more expensive than transferring control from a hardware block to a hardware block; P. 147, Sec. of "Constrained Bipartitioning Algorithms" – cost-driven constrained bi-partitioning; weight-driven constrained bi-partitioning; P. 148, R-Col., 1st Par.; Fig. 3 – proposed constrained bi-partitioning algorithms).

Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention was made to combine the teachings of Rakhmatov into the Lakshmanamurthy's system to further provide transforming the sequential application program comprises constructing a flow network model for the sequential application program; selecting a plurality of preliminary pipeline stages from the flow network model; and modifying the preliminary pipeline stages to perform control flow and

variable transmission therebetween to form the D-pipeline stages in Lakshmanamurthy system.

The motivation is that it would further enhance the Lakshmanamurthy's system by taking, advancing and/or incorporating Rakhmatov's system which offers significant advantages for providing an efficient bi-partitioning algorithm that finds an optimal solution for energy-delay product minimization and systematically searches for the best in polynomially bounded set of good solutions for delay-constrained energy minimization, and energy-constrained delay minimization, the formulation also including costs and weights as design parameters as once suggested by Rakhmatov (e.g., Sec. of "CONCLUSION").

13. **As to claim 32** (incorporating the rejection in claim 31) and **claim 35** (incorporating the rejection in claim 34), Lakshmanamurthy discloses network processor performance analysis methodology (Sec. of "ABSTRACT", 3rd Par.), but does not explicitly disclose the compiler to cause construction of a flow network model for the sequential application program, to cause selection of a plurality of preliminary pipeline stages from the flow network model and to cause modification of the preliminary pipeline stages to perform control flow and variable transformation therebetween to form the D-pipeline stages.

However, in an analogous art of hardware-software bi-partitioning for dynamically reconfigurable systems, Rakhmatov discloses the compiler to cause construction of a flow network model for the sequential application program, to cause selection of a

plurality of preliminary pipeline stages from the flow network model and to cause modification of the preliminary pipeline stages to perform control flow and variable transformation therebetween to form the D-pipeline stages (Sec. of "ABSTRACT" – a method for mapping nodes of an application control flow graph either to software or reconfigurable hardware, explicitly targeting minimization of the energy-delay cost due to both computation and configuration; using network flow techniques, after transforming the original control flow graph into an equivalent network; P. 145, R-Col., 1st Par. Through 3rd Par. – the software can directly configure the hardware, which is partially reconfigurable; partial reconfiguration allows for a selective change of hardware segments of arbitrary size at an arbitrary location, without disrupting the operation of the rest of the hardware space; such a capability greatly reduces reconfiguration time and energy consumption, because the hardware updates are highly localized. Three types of problems: (1) energy-delay product minimization, (2) energy minimization under the delay constraint, and (3) delay minimization under the energy constraint can be resolved via using network flow techniques; specifically, the cost of a node depends whether it is in software or in hardware and the cost of an edge depends whether its origin node is in software or in hardware and whether its destination node is in software or in hardware; P. 146, L-Col., 2nd Par. – 4th Par. – the cost/weight can be either the energy or the delay or the energy-delay product of a node/edge, weighted by its execution frequency; first, transferring control from a hardware block to a software block is more expensive than transferring control from a software block to a software block; second, transferring control from a software block to a hardware block is more

expensive than transferring control from a hardware block to a hardware block; P. 147, Sec. of "Constrained Bipartitioning Algorithms" – cost-driven constrained bi-partitioning; weight-driven constrained bi-partitioning; P. 148, R-Col., 1st Par.; Fig. 3 – proposed constrained bi-partitioning algorithms).

Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention was made to combine the teachings of Rakhmatov into the Lakshmanamurthy's system to further provide the compiler to cause construction of a flow network model for the sequential application program, to cause selection of a plurality of preliminary pipeline stages from the flow network model and to cause modification of the preliminary pipeline stages to perform control flow and variable transformation therebetween to form the D-pipeline stages in Lakshmanamurthy system.

The motivation is that it would further enhance the Lakshmanamurthy's system by taking, advancing and/or incorporating Rakhmatov's system which offers significant advantages for providing an efficient bi-partitioning algorithm that finds an optimal solution for energy-delay product minimization and systematically searches for the best in polynomially bounded set of good solutions for delay-constrained energy minimization, and energy-constrained delay minimization, the formulation also including costs and weights as design parameters as once suggested by Rakhmatov (e.g., Sec. of "CONCLUSION").

14. Claims 3-4, 8, 10, 13-14, 18, 20, 33, and 36 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lakshmanamurthy in view of Rakhmatov and further in view of Robschink et al., ("*Efficient Path Conditions in Dependence Graphs", May 2002, ACM*) (hereinafter 'Robschink')

15. **As to claim 3** (incorporating the rejection in claim 2) and **claim 13** (incorporating the rejection in claim 12), Rakhmatov discloses constructing the flow network model (Sec. of "ABSTRACT" – a method for mapping nodes of an application control flow graph either to software or reconfigurable hardware, explicitly targeting minimization of energy-delay cost due to both computation and configuration; show how these problems can be tackled by using network flow techniques, after transforming the original control flow graph into an equivalent network), but Lakshmanamurthy and Rakhmatov do not explicitly disclose transforming the application program into a static, single-assignment form; building a control flow graph for a loop body of the application program; building a dependence graph based on a summary graph of the control flow graph and identified, strongly-connected components (SSC) of the control flow graph; and constructing the flow network model according to a summary graph of the dependence graph and identified SSC nodes of the dependence graph.

However, in an analogous art of efficient path conditions in dependence graphs, Robschink discloses transforming the application program into a static, single-assignment form (Sec. 2.2 – Path conditions, 2nd Pa. – since there may be assignments to the same variable at different program points, all programs must be transformed into

static single assignment form (SSA) first. In SSA from, there is at most one assignment to every variable. If necessary, we will distinguish different SSA-variants of a program variable by additional indices; P. 480, 3rd Par. – since the program is transformed to SSA form first, some additional constraints must be generated which represent the Φ-function occurring in SSA form); building a control flow graph for a loop body of the application program; building a dependence graph based on a summary graph of the control flow graph (Sec. 1 – Introduction, 4th Par. – *ValSoft* can build a dependence graph for 50000 lines of C; forward and backward slices or chops can be interactively computed and visualized in the source text; Fig. 1 – a *mergesort* program and part of its SDG (System Dependence Graph); Sec. 2.1 – Dependence and slices, 2nd Par. – slices can be defined via the system dependence graph (SDG); Sec. 3 – Basic Analysis, 1st Par.; Sec. 3.1 – Analyzing data flow, 1st Par.) and identified, strongly-connected components (SSC) of the control flow graph; and constructing the flow network model according to a summary graph of the dependence graph and identified SSC nodes of the dependence graph (Sec. 4.2 – Exploiting interval analysis, 2nd Par through 3rd Par.).

Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention was made to combine the teachings of Robschink into the Lakshmanamurthy-Rakhmatov's system to further provide transforming the application program into a static, single-assignment form; building a control flow graph for a loop body of the application program; building a dependence graph based on a summary graph of the control flow graph and identified, strongly-connected components (SSC) of the control flow graph; and constructing the flow network model according to a summary

graph of the dependence graph and identified SSC nodes of the dependence graph in Lakshmanamurthy-Rakhmatov system.

The motivation is that it would further enhance the Lakshmanamurthy-Rakhmatov's system by taking, advancing and/or incorporating Robschink's system which offers that path conditions in dependence graphs are a valuable tool for various kinds of program analysis, such as program understanding or safety checks as once suggested by Robschink (Sec. of "CONCLUSION AND FUTURE WORK", 1st Par.).

16. **As to claim 8** (incorporating the rejection in claim 2) and **claim 18** (incorporating the rejection in claim 12), Rakhmatov discloses selecting the plurality of preliminary pipeline stages comprises cutting the flow network model into D-1 successive cuts, such that each cut is a balanced minimum cost cut (Sec. 1 – Introduction, 4th Par. – cost function involve costs of all nodes and all edges in the CFG, and not just the edges in the cut-set separating software-mapped nodes and hardware-mapped nodes; specifically, the cost of a node depends whether it is in software or in hardware, and the cost of an edge depends whether its origin node is in software or in hardware and whether its destination node is in software or in hardware; Sec. 3 – Related Work, 2nd Par. (Circuit Partitioning) Through 3rd Par., 6th Par. (Our Contribution) – our contribution is to show how a CFG (Control Flow Graph) with node and edge costs can be transformed into a network, so that a minimum cut in the network corresponds to an optimal bi-partition of the CFG; P. 147, 1st Par. through 2nd Par.; Fig. 2 – unconstrained bi-partitioning algorithm – *CUT = FindMinCut(V,E)*).

17. **As to claim 4** (incorporating the rejection in claim 3) and **claim 14** (incorporating the rejection in claim 13), Rakhmatov discloses constructing the flow network model comprises assigning a unique source node and a unique sink node to the flow network model (Sec. 4 – Proposed Solution, 1st Par.).

Robschink discloses adding a program node to the flow network model for each SSC node identified in the summary graph of the dependence graph (Sec. 2.2 – Path conditions, 2nd Pa. – since there may be assignments to the same variable at different program points, all programs must be transformed into static single assignment form (SSA) first. In SSA from, there is at most one assignment to every variable. If necessary, we will distinguish different SSA-variants of a program variable by additional indices; P. 480, 3rd Par. – since the program is transformed to SSA form first, some additional constraints must be generated which represent the Φ-function occurring in SSA form); adding a variable node to the flow network model for each variable that is defined and used by multiple program nodes; adding a control node C to the flow network model for each SSC node identified in the summary graph of the dependence graph as a source of control dependence (Sec. 1 – Introduction, 4th Par. – *ValSoft* can build a dependence graph for 50000 lines of C; forward and backward slices or chops can be interactively computed and visualized in the source text; Fig. 1 – a *mergesort* program and part of its SDG (System Dependence Graph); Sec. 2.1 – Dependence and slices, 2nd Par. – slices can be defined via the system dependence graph (SDG); Sec. 3 – Basic Analysis, 1st Par.; Sec. 3.1 – Analyzing data flow, 1st Par.).

Further, Rakhmatov discloses generating edges having an associated weight to connect corresponding program nodes to corresponding variable nodes; generating edges having an associated weight to connect corresponding program nodes to corresponding control nodes; and generating edges between the program nodes and one of the source node and the sink node (Sec. 2 – Problem Description, 1st Par. through 4th Par., and 6th Par.; P. 147, L-Col.,1st Par.; Sec. 4 – Proposed Solution, 1st Par. through 2nd Par.; P. 147, R-Col., 1st Par. through 4th Par.).

18. **As to claim 10** (incorporating the rejection in claim 2) and **claim 20** (incorporating the rejection in claim 12), Rakhmatov discloses modifying the preliminary pipeline stages comprises selecting a preliminary pipeline stage; altering the selected preliminary pipeline stage to enable proper transmission of live variables to and from the selected preliminary pipeline stage; altering the selected preliminary pipeline stage to enable proper transmission of control flow to and from the selected preliminary pipeline stage; and repeating the selecting, altering and altering for each preliminary pipeline stage to form the D-pipeline stages of a parallel network application (Sec. of "ABSTRACT" – a method for mapping nodes of an application control flow graph either to software or reconfigurable hardware, explicitly targeting minimization of the energy-delay cost due to both computation and configuration; using network flow techniques, after transforming the original control flow graph into an equivalent network; P. 145, R-Col., 1st Par. Through 3rd Par. – the software can directly configure the hardware, which is partially reconfigurable; partial reconfiguration allows for a selective change of

hardware segments of arbitrary size at an arbitrary location, without disrupting the operation of the rest of the hardware space; such a capability greatly reduces reconfiguration time and energy consumption, because the hardware updates are highly localized. Three types of problems: (1) energy-delay product minimization, (2) energy minimization under the delay constraint, and (3) delay minimization under the energy constraint can be resolved via using network flow techniques; specifically, the cost of a node depends whether it is in software or in hardware and the cost of an edge depends whether its origin node is in software or in hardware and whether its destination node is in software or in hardware; P. 146, L-Col., 2nd Par. – 4th Par. – the cost/weight can be either the energy or the delay or the energy-delay product of a node/edge, weighted by its execution frequency; first, transferring control from a hardware block to a software block is more expensive than transferring control from a software block to a software block; second, transferring control from a software block to a hardware block is more expensive than transferring control from a hardware block to a hardware block; P. 147, Sec. of "Constrained Bipartitioning Algorithms" – cost-driven constrained bi-partitioning; weight-driven constrained bi-partitioning; P. 148, R-Col., 1st Par.; Fig. 3 – proposed constrained bi-partitioning algorithms).

19. **As to claim 33** (incorporating the rejection in claim 32) and **claim 36** (incorporating the rejection in claim 35), Robschink discloses the compiler to cause D-1 successive cuts of the flow network mode, such that each cut is a balanced, minimum cost cut to form the D-preliminary pipeline stages (Sec. 1 – Introduction, 4th Par. – cost

function involve costs of all nodes and all edges in the CFG, and not just the edges in the cut-set separating software-mapped nodes and hardware-mapped nodes; specifically, the cost of a node depends whether it is in software or in hardware, and the cost of an edge depends whether its origin node is in software or in hardware and whether its destination node is in software or in hardware; Sec. 3 – Related Work, 2nd Par. (Circuit Partitioning) Through 3rd Par., 6th Par. (Our Contribution) – our contribution is to show how a CFG (Control Flow Graph) with node and edge costs can be transformed into a network, so that a minimum cut in the network corresponds to an optimal bi-partition of the CFG; P. 147, 1st Par. through 2nd Par.; Fig. 2 – unconstrained bi-partitioning algorithm – *CUT = FindMinCut(V,E)*).

20. Claims 9 and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lakshmanamurthy in view of Rakhmatov and Robschink and further in view of Goldberg et al., ("*A New Approach to the Maximum-Flow Problem", 1988, ACM*) (hereinafter 'Goldberg')

21. **As to claim 9** (incorporating the rejection in claim 8) and **claim 19** (incorporating the rejection in claim 18), Lakshmanamurthy, Rakhmatov, and Robschink do not disclose that cutting is performed using an iterative balanced to push-relabel algorithm.

However, in an analogous art of a new approach to the maximum-flow problem, Goldberg discloses cutting is performed using an iterative balanced to push-relabel algorithm (P. 922, 4th Par. – the algorithm pushes flow through the network to find a

blocking flow, which determines the acyclic network for the next phase; our algorithm maintains a pre-flow in the original network and pushes local flow excess toward the sink along what it estimates to be shortest paths in the residual graph; P. 924, 4th Par. – the pre-flow algorithm works by examining vertices other than *s* and *t* with positive flow excess and pushing excess form them to vertices estimated to be closer to the sink *t*, with the goal of getting as much excess as possible to *t*.).

Therefore, it would have been obvious to one of ordinary skill in the art, at the time the invention was made to combine the teachings of Goldberg into the Lakshmanamurthy-Rakhmatov-Robschink's system to further provide that cutting is performed using an iterative balanced to push-relabel algorithm in Lakshmanamurthy-Rakhmatov-Robschink system.

The motivation is that it would further enhance the Lakshmanamurthy-Rakhmatov- Robschink's system by taking, advancing and/or incorporating Goldberg's system which offers significant advantages that the method maintains a pre-flow in the original network and pushes local flow excess toward the sink along what are estimated to be shortest paths as once suggested by Goldberg (e.g., P. 921, 1st Par.).

Allowable Subject Matter

22. Claims 5-7, 15-17, 22-25 and 27-30 are objected to as being dependent upon a rejected base claim, but would be allowable if rewritten to overcome the rejections under 35 U.S.C. 102(b) and/or 35 U.S.C. 103(a) set forth in this office action and to include all the limitations of the base claim and any intervening claims.

The following is an examiner's statement of reasons for allowance:

Regarding claims 5-7, 15-17, 22-25, and 27-30, prior art of record fails to reasonably show or suggest the specific edge generations having associated weights, transformation of the preliminary application program stage, and transformation of the control flow as claimed. Specifically, the methods to generate edges having an associated weight to connect corresponding program nodes to (1) corresponding variable nodes, (2) corresponding controls nodes; the method to generate the edges between program nodes and one of the source node and the sink nodes; transformation of the preliminary application program stages; and transformation of the control flow in details.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for Allowance."

Conclusion

23. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

- Archambault et al., Loop Allocation for Optimizing Compilers (Pat. No. US 6,651,246 B1)
- Lyuh et al., "High-Level Synthesis for Low Power Based on Network Flow Method", Jun. 2003, IEEE, pp. 364-375
- Yang et al., "Efficient Network Flow Based Min-Cut Balanced Partition", 1994, ACM, pp. 50-55

24. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Ben C. Wang whose telephone number is 571-270-1240. The examiner can normally be reached on Monday - Friday, 8:00 a.m. - 5:00 p.m., EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Tuan Q. Dam can be reached on 571-272-3695. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

TUAN DAM
SUPERVISORY PATENT EXAMINER

BCW

June 6, 2007